

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-222719
(43)Date of publication of application : 17.08.2001

(51)Int.CI. G06T 7/60
G06T 1/00
G06T 7/00
G06T 7/20

(21)Application number : 2000-363449 (71)Applicant : MATSUSHITA ELECTRIC IND CO LTD
(22)Date of filing : 29.11.2000 (72)Inventor : MATSUO HIDEAKI
IMAGAWA KAZUYUKI
TAKADA YUJI
EJIMA TOSHIAKI
BABA KOJUN

(30)Priority

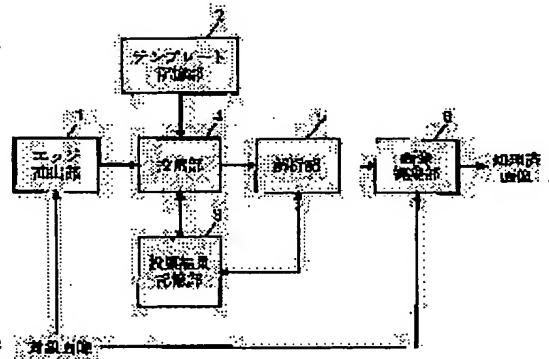
Priority number : 11342025 Priority date : 01.12.1999 Priority country : JP

(54) FACE EXTRACTING DEVICE, FACE EXTRACTING METHOD AND RECORDING MEDIUM FOR FACE EXTRACTION PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a face extracting device and a face extracting method which can speedily extract the position and size of a face from various object images and have high versatility.

SOLUTION: An edge extracting part 1 extracts an edge part from an object image and generates an edge image. A template storing part 2 stores a plurality of templates with central points provided concentrically in various sizes obtained by making a previously defined shape different in size by analogy. A vote result storing part 3 is provided with a vote storage area where the results of vote processing performed in a voting part 4 in each size shape constituting a template. The voting part 4 performs vote processing by using the templates at each pixel position of the edge image and stores the vote results in each vote storage area. An analyzing part 5 evaluates a cluster on the basis of the vote results stored in the vote storage area and calculates the position and the size of a face included in the object image after the vote processing is completed in the part 4.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] A face extractor which asks for a location and magnitude of a face which are characterized by providing the following, and which are contained in an object image. The edge extract section which extracts the edge section (pixel equivalent to a person's outline, an outline of a face, etc.) from said object image, and generates an image (henceforth an edge image) of only the edge section concerned. The template storage section which memorizes a template which established two or more configurations defined beforehand in the shape of the said heart in the central point with various sizes which changed magnitude by similarity. The vote result storage section which matches a coordinate location and the number of votes on said edge image, and is memorized for every configuration of each size which constitutes said template, respectively. Each pixel location of said edge section is made to carry out sequential migration of the central point of said template. The vote section which increases or decreases said number of votes memorized by said vote result storage section, respectively about each coordinate location corresponding to a location of all pixels that form a configuration of each of said size for every pixel location concerned to which it was made to move. The analysis section which asks for a location and magnitude of a face which are contained in said object image based on said each number of votes memorized by said vote result storage section.

[Claim 2] Said configuration defined beforehand is a face extractor according to claim 1 characterized by being a circle.

[Claim 3] Said configuration defined beforehand is a face extractor according to claim 1 characterized by being an ellipse.

[Claim 4] Said configuration defined beforehand is a face extractor according to claim 1 characterized by being a polygon.

[Claim 5] Said edge extract section is a face extractor according to claim 1 to 4 characterized by extracting said edge section by filtering in said object image and asking for a high frequency component.

[Claim 6] Said edge extract section is a face extractor according to claim 1 to 4 characterized by extracting said edge section by asking for difference with an image which constitutes said object image and which gets mixed up on the present image and a time amount target for every image.

[Claim 7] Said edge extract section is a face extractor according to claim 5 or 6 which detects a pixel which exists in a leftmost side of each scan line in the field concerned, and right-hand side for every predetermined rectangle field among extracted pixels, and is characterized by making only the detected pixel concerned into said edge section.

[Claim 8] Said analysis section is a face extractor according to claim 1 to 7 which clusters to said each number of votes memorized by said vote result storage section, and is characterized by performing a location of a face contained in said object image, and narrowing down of magnitude.

[Claim 9] A face extractor according to claim 1 to 8 which distinguishes a face field appointed with a location and magnitude of a face which were called for in said analysis section, and the other field, and is further equipped with the image editorial department which performs predetermined edit to said object image.

[Claim 10] A face extractor according to claim 1 to 8 further equipped with the image editorial department which performs edit which replaces an image of a face field appointed with a location

and magnitude of a face which were called for in said analysis section by other images.

[Claim 11] A face extract method of asking for a location and magnitude of a face which are characterized by providing the following and which are contained in an object image An extract step which extracts the edge section (pixel equivalent to a person's outline, an outline of a face, etc.) from said object image, and generates an image (henceforth an edge image) of only the edge section concerned The 1st storage step which memorizes a template which established two or more configurations defined beforehand in the shape of the said heart in the central point with various sizes which changed magnitude by similarity The 2nd storage step which matches a coordinate location and the number of votes on said edge image, and is memorized for every configuration of each size which constitutes said template, respectively Each pixel location of said edge section is made to carry out sequential migration of the central point of said template. A vote step which increases or decreases said number of votes memorized at said 2nd storage step, respectively about each coordinate location corresponding to a location of all pixels that form a configuration of each of said size for every pixel location concerned to which it was made to move, An analysis step which asks for a location and magnitude of a face which are contained in said object image based on said each number of votes after processing of said vote step

[Claim 12] Said configuration defined beforehand is the face extract method according to claim 11 characterized by being a circle.

[Claim 13] Said configuration defined beforehand is the face extract method according to claim 11 characterized by being an ellipse.

[Claim 14] Said configuration defined beforehand is the face extract method according to claim 11 characterized by being a polygon.

[Claim 15] Said extract step is the face extract method according to claim 11 to 14 characterized by extracting said edge section by filtering in said object image and asking for a high frequency component.

[Claim 16] Said extract step is the face extract method according to claim 11 to 14 characterized by extracting said edge section by asking for difference with an image which constitutes said object image and which gets mixed up on the present image and a time amount target for every image.

[Claim 17] Said extract step is the face extract method according to claim 15 or 16 which detects a pixel which exists in a leftmost side of each scan line in the field concerned, and right-hand side for every predetermined rectangle field among extracted pixels, and is characterized by making only the detected pixel concerned into said edge section.

[Claim 18] Said analysis step is the face extract method according to claim 11 to 17 which clusters to said each number of votes after processing of said vote step, and is characterized by performing a location of a face contained in said object image, and narrowing down of magnitude.

[Claim 19] A face extract method of asking for a location and magnitude of a face which are contained in an object image is data medium recorded as a program which can be performed on a computer apparatus. The edge section (pixel equivalent to a person's outline, an outline of a face, etc.) is extracted from said object image. A configuration beforehand determined as an extract step which generates an image (henceforth an edge image) of only the edge section concerned with various sizes which changed magnitude by similarity The 1st storage step which memorizes a template prepared in the shape of the said heart in the central point, [two or more] The 2nd storage step which matches a coordinate location and the number of votes on said edge image, and is memorized for every configuration of each size which constitutes said template, respectively, Each pixel location of said edge section is made to carry out sequential migration of the central point of said template. A vote step which increases or decreases said number of votes memorized at said 2nd storage step, respectively about each coordinate location corresponding to a location of all pixels that form a configuration of each of said size for every pixel location concerned to which it was made to move, A record medium which recorded a program for performing at least an analysis step which asks for a location and magnitude of a face which are contained in said object image based on said each number of votes after processing of said vote step.

[Claim 20] Said configuration defined beforehand is a record medium according to claim 19 characterized by being a circle.

[Claim 21] Said configuration defined beforehand is a record medium according to claim 19

characterized by being an ellipse.

[Claim 22] Said configuration defined beforehand is a record medium according to claim 19 characterized by being a polygon.

[Claim 23] Said extract step is a record medium according to claim 19 to 22 characterized by extracting said edge section by filtering in said object image and asking for a high frequency component.

[Claim 24] Said extract step is a record medium according to claim 19 to 22 characterized by extracting said edge section by asking for difference with an image which constitutes said object image and which gets mixed up on the present image and a time amount target for every image.

[Claim 25] Said extract step is a record medium according to claim 23 or 24 which detects a pixel which exists in a leftmost side of each scan line in the field concerned, and right-hand side for every predetermined rectangle field among extracted pixels, and is characterized by making only the detected pixel concerned into said edge section.

[Claim 26] Said analysis step is a record medium according to claim 19 to 25 which clusters to said each number of votes after processing of said vote step, and is characterized by performing a location of a face contained in said object image, and narrowing down of magnitude.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any
damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] this invention is technology more specifically used for an image processing about the record medium of a face extract program at a face extractor and the face extract method list, and relates to data medium which recorded the program for performing the method concerned on the equipment and the method list which use the template and extract the location and the magnitude of a face which are contained in an object image at a high speed.

[0002]

[Description of the Prior Art] Needless to say, a face has the important semantics of expressing thinking of a person and feeling. therefore, in the field of an image processing which deals with an image including a person, if there is a system by which a person's face can detect and process automatically whether it exists in the magnitude which is how much in which location of an object image about a static image or a dynamic image (a person also contains the image artificially generated with computer graphics besides the actually photoed image etc.), it is convenient. For this reason, the attempt which is going to extract the field of a face from the target image on an image processing system has started in recent years.

[0003] As a Prior art which extracts the field of a face from such an image, what is indicated by JP,9-73544,A (henceforth the 1st reference), JP,10-307923,A (henceforth the 2nd reference), etc. exists. It is performing the technology which approximates a face field with an ellipse being indicated by this 1st reference, changing serially five parameters (the ratio b of a main coordinate (x y), a radius r, a major axis, and a minor axis, and the angle theta of a major axis and a x axis to make) which specify an ellipse into it, and calculating the optimal value for the extract of a face field. Moreover, the technology which finds the components (an eye, a nose, opening, etc.) of a face in the 2nd reference in order, and goes to it is indicated.

[0004]

[Problem(s) to be Solved by the Invention] However, since the Prior art indicated by the 1st reference of the above will estimate approximation, changing each parameter, huge repeat count is needed (it is because processing of angle theta is heavy especially). For this reason, it is hopeless to make the changing portrait image follow in the throughput of the present personal computer level using this Prior art, and to perform face extract processing in the real time. Moreover, in this Prior art, when two or more persons' face is contained in one image, there is no consideration of, and it must be said that versatility is low. Moreover, in the Prior art indicated by the 2nd reference of the above, probably, if the location of a face field is not known, extract processing of the components of a face cannot be performed. Therefore, in this Prior art, it can apply only to a specific image but versatility is low.

[0005] So, the purpose of this invention is providing the face extractor with high versatility and the face extract method list which can extract the location and magnitude of a face from various object images at a high speed with data medium which recorded the program for performing the method concerned.

[0006]

[The means for solving a technical problem and an effect of the invention] The 1st invention is a face extractor which asks for the location and magnitude of a face which are contained in an object

image, and extracts the edge section (pixel equivalent to a person's outline, the outline of a face, etc.) from an object image. The configuration beforehand determined as the edge extract section which generates the image (edge image) of only the edge section concerned with the various sizes which changed magnitude by similarity. The template storage section which memorizes the template prepared in the shape of the said heart in the central point, [two or more] The vote result storage section which matches the coordinate location and the number of votes on an edge image, and is memorized for every configuration of each size which constitutes a template, respectively, For every pixel location concerned which made each pixel location of the edge section carry out sequential migration, and moved the central point of a template to it. The vote section which increases or decreases the number of votes memorized by the vote result storage section, respectively about each coordinate location corresponding to the location of all the pixels that form the configuration of each size, Based on each number of votes memorized by the vote result storage section, it has the analysis section which asks for the location and magnitude of a face which are contained in an object image. As mentioned above, according to the 1st invention, the location of a face is detectable at a high speed only by the vote processing with a light processing burden, and its evaluation. And since the template equipped with the configuration of said heart-like two or more sizes by similarity is used, the edge section which probably includes the face field will perform substantial approximation whether to be close to which size of these configurations, and can also extract the magnitude of a face at a high speed. Thus, by 1st invention, since a processing burden is sharply mitigable, the throughput of the present personal computer level can also extract a face field in the real time mostly. Moreover, since the number of the point in which portion a face field is among object images in the 1st invention, or a face etc. is unknown, does not interfere before extract processing and can detect a face field uniformly about various object images of a large range, its versatility is very high.

[0007] The 2nd – the 4th invention are invention which is subordinate to the 1st invention, respectively, and the configuration defined beforehand is characterized by being either a circle, an ellipse or a polygon. As mentioned above, the 2nd – the 4th invention show the typical configuration used for a template. Precision of a vote result can be made high by using for the configuration of a template a circle with a fixed distance to each pixel which forms a configuration from the central point especially.

[0008] The 5th invention is invention subordinate to the 1st – the 4th invention, and the edge extract section is characterized by extracting the edge section by filtering in an object image and asking for a high frequency component. As mentioned above, since according to the 5th invention it filters in an object image and asks for a high frequency component, it is suitable for the location of a face in case an object image is a static image, and detection of magnitude.

[0009] The 6th invention is invention subordinate to the 1st – the 4th invention, and the edge extract section is characterized by extracting the edge section by asking for difference with the image which constitutes an object image and which gets mixed up on the present image and a time amount target for every image. As mentioned above, since it asks for the difference of a current object image and the image which gets mixed up in time in it according to the 6th invention, it is suitable for the location of the face in each coma in case an object image is a dynamic image, and detection of magnitude. Moreover, since the template is used for detection, like [at the time of zoom-in and a barn], it is stabilized even when change of a face is large, and a face field can be extracted at a high speed.

[0010] The 7th invention is invention subordinate to the 5th and 6th invention, and among the extracted pixels, for every predetermined rectangle field, the edge extract section detects the pixel which exists in the leftmost side of each scan line in the field concerned, and right-hand side, and is characterized by making only the detected pixel concerned into the edge section. As mentioned above, according to the 7th invention, extract processing of a face field can be performed at a high speed by that which can obtain the edge section which eliminated the portion which the difference of an inside texture produces, and which consists only of a person's outline.

[0011] The 8th invention is invention subordinate to the 1st – the 7th invention, and the analysis section clusters to each number of votes memorized by the vote result storage section, and is characterized by performing the location of the face contained in an object image, and narrowing

down of magnitude. As mentioned above, even when two or more faces exist in an object image by dealing with a vote result (each number of votes) as a cluster according to the 8th invention, the mutual relevance of a vote result can be evaluated justly and a face field can be extracted.

[0012] The 9th invention is invention subordinate to the 1st – the 8th invention, distinguishes the face field appointed with the location and magnitude of a face which were called for in the analysis section, and the other field, and is further equipped with the image editorial department which performs predetermined edit to an object image. As mentioned above, according to the 9th invention, by distinguishing and editing the face field appointed with the location and magnitude of a face of an object image, and the other field, only a face can be emphasized and an image with the legible portion for which it asks can be obtained. For example, portions other than a face field can be made solid coating by a fixed texture and a fixed color, and an interesting effect can be aimed at.

[0013] The 10th invention is invention subordinate to the 1st – the 8th invention, and is further equipped with the image editorial department which performs edit which replaces the image of the face field appointed with the location and magnitude of a face which were called for in the analysis section by other images. As mentioned above, according to the 10th invention, a face can be intentionally hidden by replacing the image of the face field appointed with the location and magnitude of a face of an object image by other images. For example, like [in the case of using for the dynamic image which supervises demented elderly's wandering], although it is necessary to fully supervise, when privacy should be protected, the existence of a face itself can replace the portion of a face by another image, and it can respond. Furthermore, also when picturizing actuation of human being and transposing this to other characters, it can apply.

[0014] The 11th invention is the face extract method of asking for the location and magnitude of a face which are contained in an object image, and extracts the edge section from an object image. The configuration beforehand determined as the extract step which generates an edge image with the various sizes which changed magnitude by similarity The 1st storage step which memorizes the template prepared in the shape of the said heart in the central-point, [two or more] The 2nd storage step which matches the coordinate location and the number of votes on an edge image, and is memorized for every configuration of each size which constitutes a template, respectively, For every pixel location concerned which made each pixel location of the edge section carry out sequential migration, and moved the central point of a template to it The vote step which increases or decreases the number of votes memorized at the 2nd storage step, respectively about each coordinate location corresponding to the location of all the pixels that form the configuration of each size, Based on each number of votes after processing of a vote step, it has the analysis step which asks for the location and magnitude of a face which are contained in an object image. As mentioned above, according to the 11th invention, the location of a face is detectable at a high speed only by the vote processing with a light processing burden, and its evaluation. And since the template equipped with the configuration of said heart-like two or more sizes by similarity is used, the edge section which probably includes the face field will perform substantial approximation whether to be close to which size of these configurations, and can also extract the magnitude of a face at a high speed. Thus, by 11th invention, since a processing burden is sharply mitigable, the throughput of the present personal computer level can also extract a face field in the real time mostly. Moreover, since the number of the point in which portion a face field is among object images in the 11th invention, or a face etc. is unknown, does not interfere before extract processing and can detect a face uniformly about various object images of a large range, its versatility is very high.

[0015] The 12th – the 14th invention are invention which is subordinate to the 11th invention, respectively, and the configuration defined beforehand is characterized by being either a circle, an ellipse or a polygon. As mentioned above, the 12th – the 14th invention show the typical configuration used for a template. Precision of a vote result can be made high by using for the configuration of a template a circle with a fixed distance to each pixel which forms a configuration from the central point especially.

[0016] The 15th invention is invention subordinate to the 11th – the 14th invention, and an extract step is characterized by extracting the edge section by filtering in an object image and asking for a high frequency component. As mentioned above, since according to the 15th invention it filters in an

object image and asks for a high frequency component, it is suitable for the location of a face in case an object image is a static image, and detection of magnitude.

[0017] The 16th invention is invention subordinate to the 11th – the 14th invention, and an extract step is characterized by extracting the edge section by asking for difference with the image which constitutes an object image and which gets mixed up on the present image and a time amount target for every image. As mentioned above, since it asks for the difference of a current object image and the image which gets mixed up in time in it according to the 16th invention, it is suitable for the location of the face in each coma in case an object image is a dynamic image, and detection of magnitude. Moreover, since the template is used for detection, like [at the time of zoom-in and a barn], it is stabilized even when change of a face is large, and a face field can be extracted at a high speed.

[0018] The 17th invention is invention subordinate to the 15th and 16th invention, and among the extracted pixels, for every predetermined rectangle field, an extract step detects the pixel which exists in the leftmost side of each scan line in the field concerned, and right-hand side, and is characterized by making only the detected pixel concerned into the edge section. As mentioned above, according to the 17th invention, extract processing of a face field can be performed at a high speed by that which can obtain the edge section which eliminated the portion which the difference of an inside texture produces, and which consists only of a person's outline.

[0019] The 18th invention is invention subordinate to the 11th – the 17th invention, and an analysis step clusters to each number of votes after processing of a vote step, and is characterized by performing the location of the face contained in an object image, and narrowing down of magnitude. As mentioned above, even when two or more faces exist in an object image by dealing with a vote result (each number of votes) as a cluster according to the 18th invention, the mutual relevance of a vote result can be evaluated justly and a face field can be extracted.

[0020] The face extract method of asking for the location and magnitude of a face which are contained in an object image is data medium recorded as a program which can be performed on a computer apparatus, and the 19th invention extracts the edge section from an object image. The configuration beforehand determined as the extract step which generates an edge image with the various sizes which changed magnitude by similarity The 1st storage step which memorizes the template prepared in the shape of the said heart in the central point, [two or more] The 2nd storage step which matches the coordinate location and the number of votes on an edge image, and is memorized for every configuration of each size which constitutes a template, respectively, For every pixel location concerned which made each pixel location of the edge section carry out sequential migration, and moved the central point of a template to it The vote step which increases or decreases the number of votes memorized at the 2nd storage step, respectively about each coordinate location corresponding to the location of all the pixels that form the configuration of each size, The program for performing at least the analysis step which asks for the location and magnitude of a face which are contained in an object image based on each number of votes after processing of a vote step is recorded.

[0021] The 20th – the 22nd invention are invention which is subordinate to the 19th invention, respectively, and the configuration defined beforehand is characterized by being either a circle, an ellipse or a polygon.

[0022] The 23rd invention is invention subordinate to the 19th – the 22nd invention, and an extract step is characterized by extracting the edge section by filtering in an object image and asking for a high frequency component.

[0023] The 24th invention is invention subordinate to the 19th – the 22nd invention, and an extract step is characterized by extracting the edge section by asking for difference with the image which constitutes an object image and which gets mixed up on the present image and a time amount target for every image.

[0024] The 25th invention is invention subordinate to the 23rd and 24th invention, and among the extracted pixels, for every predetermined rectangle field, an extract step detects the pixel which exists in the leftmost side of each scan line in the field concerned, and right-hand side, and is characterized by making only the detected pixel concerned into the edge section.

[0025] The 26th invention is invention subordinate to the 19th – the 25th invention, and an analysis

step clusters to each number of votes after processing of a vote step, and is characterized by performing the location of the face contained in an object image, and narrowing down of magnitude.

[0026]

[Embodiment of the Invention] Drawing 1 is the block diagram showing the configuration of the face extractor concerning 1 operation gestalt of this invention. In drawing 1, the face extractor concerning this operation gestalt is equipped with the edge extract section 1, the template storage section 2, the vote result storage section 3, the vote section 4, the analysis section 5, and the image editorial department 6. Hereafter, the actuation of each configuration of a face extractor and the face extract method concerning this operation gestalt are explained, referring to a drawing.

[0027] The edge extract section 1 inputs the image (henceforth an object image) set as the object of face extract processing, extracts the edge section from this object image, and generates the image (henceforth an edge image) of only the edge section. Here, the edge section is a portion (pixel) equivalent to a person's outline, the outline of a face, etc., and is a portion used as the high frequency component in an object image. Any of a static image or a dynamic image are sufficient as the object image inputted, and the edge extract section 1 can extract the edge section from an object image using the following technique, respectively.

[0028] In using a static image as an object image, as shown in drawing 2 (a), it uses the filter 11 which takes out a high frequency component as the edge extract section 1. Thereby, the edge section can be extracted only by letting an object image pass in a filter 11. As this filter 11, a Sobel filter is used suitably. the difference which extracts the difference (data for every pixel difference) of the present image and the image which gets mixed up on a time amount target as the edge extract section 1 as shown in drawing 2 (b) in using a dynamic image as an object image — the extract section 12 is used. It can ask for difference by this for every image which constitutes a dynamic image, respectively, and this difference can extract a large portion (portion which the motion produced) as the edge section.

[0029] In addition, by the technique mentioned above, the portion which the difference of a texture produces in the inside with a person's outline, the outline of a face, etc. is also extracted as the edge section. An example of the edge image with which such the edge section was extracted is shown in drawing 3 (a). In the face extractor of this invention, although it can process satisfactorily, in order to enable it to process at a high speed more even if the portion which the difference of a texture produces is extracted as the edge section, it is desirable to use the following technique. First, in an edge image like drawing 3 (a), the range of the edge section which concentrates to some extent and exists is surrounded in a rectangle field, respectively (drawing 3 (b)). Next, a scan is performed from the both sides of a left end and a right end in each of this rectangle field, respectively, and only the edge section detected first is determined as the edge section of an object image (drawing 3 (b)). By performing this processing about all the scan lines in a rectangle field, as shown in drawing 3 (c), the edge section which eliminated the portion which the difference of an inside texture produces and which consists only of a person's outline can be obtained. the filter 11 which mentioned above the configuration which performs such processing, or difference — what is necessary is to just be inserted in the latter part of the extract section 12

[0030] The data of the template which established two or more configurations defined beforehand in the shape of the said heart in the central point with the various sizes which changed magnitude by similarity is memorized by the template storage section 2. Although a circle, an ellipse, a regular polygon, a polygon, etc. can be used for the configuration of this template, it is most desirable to use a circle with an always fixed distance from the central point to a configuration line (each pixel which forms a configuration). Thereby, precision of the vote result mentioned later can be made high. Hereafter, this operation gestalt explains the case where the central point uses the template which prepared two or more concentric circles from which a radius differs by P, as shown in drawing 4. Two or more circles t1-tn (n is the integer of arbitration) which constitute a template here may be the configurations that a radius changes at a fixed gap, like the template T1 shown in drawing 4 (a), and may be the configurations that a radius changes at intervals of an indeterminate, like the template T2 shown in drawing 4 (b). Moreover, a part or all line breadth may consist of 2 or more (namely, circular ring configuration) dots like template T3 which all line breadth may be constituted from 1 dot (equivalent to 1 pixel of an object image), and shows to drawing 4 (c) like the template T2

which shows two or more circles t_1-t_n which constitute a template to drawing 4 (b). In addition, in the following explanation, a circle and a circular ring are named generically and it is only called a "circle."

[0031] Although it is collectively treated as one template by two or more of these circles t_1-t_n and the template storage section 2 memorizes, in actual processing, each circles t_1-t_n which constitute a template will be treated independently. For this reason, the pixel data which forms each circles t_1-t_n is memorized in table format in the template storage section 2, respectively.

[0032] The field (henceforth a vote storage region) which memorizes the result of the vote processing performed in the vote section 4 mentioned later is established in the vote result storage section 3 for every configuration of each size which constitutes the template memorized by the template storage section 2. In this example, since the configurations of each size are Circles t_1-t_n , n vote storage regions will be established in the vote result storage section 3 about Circles t_1-t_n . In addition, this vote storage region has a range corresponding to an object image.

[0033] The vote section 4 performs vote processing about the edge image generated in the edge extract section 1 using the template memorized by the template storage section 2. Drawing 5 is a flow chart which shows the procedure of the vote processing performed in the vote section 4. With reference to drawing 5, first, the vote section 4 accesses the vote result storage section 3, and initializes all the components (vote value) showing the coordinate in each vote storage region to zero (step S11). Next, the vote section 4 sets the central point P of a template to the head pixel location of the edge section in an edge image (step S12). From the upper left, this head pixel location is sequentially scanned to the upper right or the lower left, goes to it, and should just make for example, an edge image top the location of the pixel of the edge section detected first.

[0034] Next, the vote section 4 initializes to "1" the counter i which specifies the configuration (this example circles t_1-t_n) which constitutes a template (step S13). Next, the vote section 4 acquires xy coordinate on the edge image of all the pixels that form a circle t_1 about the circle t_1 specified by Counter i (= 1), respectively (step S14). And the vote section 4 casts its vote for the component showing each acquired xy coordinate by adding "1", respectively in the vote storage region about the circle t_1 prepared in the vote result storage section 3 (step S15). After this processing finishes, the vote section 4 increments one counter i, and is set to $i=2$ (step S17). Next, the vote section 4 acquires xy coordinate on the edge image of all the pixels that form a circle t_2 about the circle t_2 specified by Counter i (= 2), respectively (step S14). And the vote section 4 casts its vote for the component showing each acquired xy coordinate by adding "1", respectively in the vote storage region about the circle t_2 prepared in the vote result storage section 3 (step S15).

[0035] Similarly hereafter, incrementing every one counter i until it becomes $i=n$, about the circles t_3-t_n which are the total configurations which constitute a template, the vote section 4 repeats vote processing of the above-mentioned steps S14 and S15, and performs it (steps S16 and S17). By this, vote processing in a head pixel location will be performed to each of the vote storage region about each circles t_1-t_n . And it performs further that the vote section 4 sets the central point P of a template to the next pixel location of the edge section, repeats processing of the above-mentioned steps S13-S17, and performs it by a unit of 1 time to all the pixels of the edge section in an edge image (steps S18 and S19). That is, vote processing by the vote section 4 is performed so that the central point P of a template may crawl on all the pixels of the edge section.

[0036] For example, the number of votes as shown in drawing 6 is memorized in n vote storage regions established in the vote result storage section 3 by performing the above-mentioned vote processing to the edge image shown in drawing 3 (c). In addition, by drawing 6, in order to make a drawing legible, the case where vote processing is performed in some pixel locations of the edge section is shown. In drawing 6, the portion of a continuous line circle is equivalent to the coordinate component for which its vote was cast based on the configuration (circles t_1-t_n) of each size of a template in the above-mentioned step S15, and serves as a coordinate number "1." Moreover, since accumulation of each number of votes is carried out as mentioned above, the portion (- mark shows among drawing) which the continuous line circle of drawing 6 intersects expresses that the number of votes is so high that there are many crossing numbers.

[0037] Therefore, if vote processing mentioned above is performed to the edge section expressing the outline of the face approximated to a circle or an ellipse with the central point, the high number

of votes will concentrate near [the] the central point. Therefore, if the portion which a high vote value concentrates is judged, it will become possible to specify the center of a face. Moreover, the phenomenon which such a high vote value concentrates appears more notably, when the circle configuration which has a very near radius equal to the minimum width of face of the edge section expressing the outline of a face or also in a template is used. Therefore, if this phenomenon judges whether it has appeared in the vote storage region of which circle configuration notably, it will become possible to specify the magnitude of a face. It can be said that this point resembles generalization Hough conversion. However, by the face extract method of this invention, it differs from generalization Hough conversion clearly in that the magnitude can be specified at once with the central point of the edge section by using the template which has the configuration of two or more sizes in the shape of the said heart.

[0038] In addition, in the above-mentioned step S11, all the components showing the coordinate in each vote storage region are initialized to the maximum defined beforehand, and its vote may be cast in the above-mentioned step S15 by subtracting "1" from the component showing each acquired xy coordinate, respectively. In this case, it is possible to specify the center of a face, if the portion which a low vote value concentrates is judged, and if it judges whether this concentration phenomenon has appeared in the vote storage region of which circle configuration notably, it will become possible to specify the magnitude of a face. Moreover, it sets to the above-mentioned step S15, and the value which makes the number of votes add or subtract may be except "1", and can set up a value freely.

[0039] Next, based on the vote result memorized by the vote result storage section 3, the technique of pinpointing the face field of an object image is explained. After the vote processing by the vote section 4 is completed, based on the vote result memorized by the vote result storage section 3, the analysis section 5 evaluates the cluster and asks for the location and magnitude of a face which are contained in an object image. Drawing 7 is a flow chart which shows the procedure of the analysis processing performed in the analysis section 5.

[0040] With reference to drawing 7, the analysis section 5 sets to "1" the counter j which specifies first the configuration (this example circles t1-tn) which constitutes a template (step S21). Next, the analysis section 5 extracts only the component exceeding threshold G (for example, 200th grade) which the number of votes defined beforehand about the circle t1 specified by Counter j (= 1) with reference to the vote result memorized in the vote storage region about the circle t1 of the vote result storage section 3 (step S22). This threshold G can be set to arbitration based on the definition and the extract precision for which it asks of an object image. Next, only for the extracted component, the analysis section 5 clusters (step S23), and calculates the variance and covariance value of each field by which clustering was carried out, respectively (step S24). The similarity in this clustering may be judged using any of Euclid square distance, standardization Euclid square distance, Mahalanobis' generalized distance, or MINKOFU skiing distance. Moreover, any of a minimum distance method (SLINK:single linkage clustering method), the longest distance method (CLINK:complete linkage clustering method), or the group method of averaging (UPGMA:unweighted pair-group method using arithmetic averages) may be used for formation of a cluster. Next, the analysis section 5 compares the variance and covariance value of each field by which clustering was carried out with threshold H defined beforehand (step S25). When each value is under threshold H in this step S25, and the analysis section 5 It considers that the central point of that field is the central point of a face, size (diameter) of a circle t1 which the counter j at this time (= 1) points out is used as the minor-axis length of a face (step S26), and the length which added constant value (it sets experientially) to this minor-axis length is determined as major-axis length of a face (step S27). And the analysis section 5 holds this determined central point, minor-axis length, and major-axis length as an analysis result (step S28). On the other hand, when each value is more than threshold H in the above-mentioned step S25, the analysis section 5 judges that the central point of the field is not the central point of a face, and it moves from it to the next processing.

[0041] After this processing finishes, the analysis section 5 increments one counter j, and is set to j= 2 (step S30). Next, the analysis section 5 extracts only the component exceeding threshold G which the number of votes defined beforehand about the circle t2 specified by Counter j (= 2) with reference to the vote result memorized in the vote storage region about the circle t2 of the vote

result storage section 3 (step S22). Next, only for the extracted component, the analysis section 5 clusters (step S23), and calculates the variance and covariance value of each field by which clustering was carried out, respectively (step S24). Next, the analysis section 5 compares the variance and covariance value of each field by which clustering was carried out with threshold H defined beforehand (step S25). And when each value is under threshold H in this step S25, the analysis section 5 considers that the central point of that field is the central point of a face, size of a circle t2 which the counter j at this time (= 2) points out is used as the minor-axis length of a face (step S26), and the length which added constant value to this minor-axis length is determined as major-axis length of a face (step S27). And the analysis section 5 adds this determined central point, minor-axis length, and major-axis length as an analysis result, and holds them (step S28). On the other hand, when each value is more than threshold H in the above-mentioned step S25, the analysis section 5 judges that the central point of the field is not the central point of a face, and it moves from it to the next processing.

[0042] Similarly hereafter, incrementing every one counter j until it becomes $j=n$, about the vote storage region about each circles t_3-t_n memorized by the vote result storage section 3, the analysis section 5 repeats analysis processing of the above-mentioned steps S22-S28, and performs it (steps S29 and S30). Thereby, the analysis result of the face field extract in the vote storage region about each circles t_1-t_n can be obtained. This analysis result is outputted to the image editorial department 6.

[0043] Here, the clustering processing performed at the above-mentioned steps S23 and S24 is briefly explained with reference to drawing 8. The component (- mark in drawing) for which the number of votes exceeds threshold G assumes the case where it is distributed like drawing 8 (a). In such a case, evaluation by the following clustering is performed in the analysis section 5. In early clustering, four initial clusters A, B, C, and D are generated like drawing 8 (b), for example. If an initial cluster is generated next, the mutual similarity of the initial clusters A, B, C, and D is called for, and if similarity is more than a predetermined threshold, association of the clusters will be performed. It is the case where Cluster C and Cluster D are combined and it has become Cluster E in the example of drawing 8 (c). And the variance of the final clusters A, B, and E etc. is calculated and evaluated, and it is considered that the central point of the clusters A and B with a small variance is the center of a face. Since a variance becomes large, it is not considered that Cluster E is the center of a face.

[0044] In addition, it is good to judge a face field by evaluation of a variance etc., based on the following point, when two or more clusters are detected. As the 1st, the central point is the same, and when two or more detection of the cluster from which magnitude differs is carried out, let the cluster which has the minimum variance among these clusters be a face field. When two or more detection of the cluster from which the central point differs and magnitude differs as the 2nd is carried out, it is judged as that in which the face field of the magnitude which is different in a different location exists. When the central point differs and two or more detection of the cluster with the same magnitude is carried out as the 3rd, it is judged as that in which the face field of the same magnitude as a different location exists.

[0045] The image editorial department 6 inputs an analysis result (face field) from the analysis section 5, and performs various processings for which it asks to an object image using a face field and the other field being distinguishable with an analysis result. For example, the image editorial department 6 can get the image which emphasized only the face by leaving only a face field for the object image shown in drawing 9 (a), as shown in drawing 9 (b), cutting off a background, or carrying out the solid coating of the background by one color or the single texture. Moreover, the image of the face field of the object image shown in drawing 9 (a) can be transposed to other images (face image of other characters etc.), as shown in drawing 9 (c). This becomes possible to hide a face intentionally. In addition, this image editorial department 6 is a configuration prepared if needed, in order to perform various processings for which it asks using the extracted face field, and it is not a configuration indispensable to the face extractor of this invention.

[0046] As mentioned above, according to the face extractor and method concerning 1 operation gestalt of this invention, a burden is only evaluation of light vote processing (fundamentally addition processing chisel) and the number of votes, and the location of a face can be detected at a high

speed. And since the template equipped with the configuration of said heart-like two or more sizes by similarity is used, the edge section which will be a face field will perform substantial approximation whether to be close to which size of these configurations, and can also extract the magnitude of a face at a high speed. Thus, by this invention, since a processing burden is sharply mitigable, the throughput of the present personal computer level can also extract a face field in the real time mostly. Moreover, since the number of the point in which portion a face field is among object images in this invention, or a face etc. is unknown, and there is no inconvenience before extract processing and it can detect a face uniformly about various object images of a large range, its versatility is very high. Even when two or more faces exist in an object image by clustering a vote result especially, the mutual relevance of a vote result can be evaluated justly and a face field can be extracted.

[0047] In addition, each function (the face extract method) which the face extractor concerning the above-mentioned operation gestalt realizes typically is realized by the storage (ROM, RAM, hard disk, etc.) with which the predetermined program was stored, and CPU (central processing unit) which performs the program concerned. In this case, each program may be introduced through record media, such as CD-ROM and a floppy (registered trademark) disk. The case where a program is dispersively recorded on two or more record media, and is distributed is included in this installation.

[0048] When a program does not ask whether they are some operating systems but is making a various process or various threads (DLL etc.) take over a part of the function here, even if it is the record medium with which the program part concerning the function made to take over is not stored, it corresponds to the "record medium" with which the program which performs the face extract method of this invention was recorded.

[0049] Moreover, the face extract method of this invention may be realized in server-client format, although the case where it realized in stand-alone format (drawing 1) was illustrated. That is, it does not interfere on a server connectable with the terminal which is a client, or a network besides the stand-alone format that all the functions of the face extract method are included only in one terminal, in the server-client format that the function of all or a part of the face extract methods is included, either. For example, most functions are given to a server side and only a WWW browser is given in a client side. In this case, although various kinds of information (a template, the number of votes, etc.) is usually on a server and it is fundamentally distributed to a client via a network, when that information is on a server, the storage of that server corresponds to a "record medium", and when that information is on a client, the recording device of that client will correspond to a "record medium."

[0050] Furthermore, the program which performs the face extract method of this invention may be which the following gestalt. The gestalt of the application which was compiled and became absolute language. The gestalt of the pseudo code interpreted by an above-mentioned process or an above-mentioned thread. The gestalt with which a resource and a source code, and the compiler and linker that can generate the application of absolute language from these are stored in a "record medium" at least. The gestalt with which the interpreter which can generate the application of a pseudo code is stored in a "record medium" from a resource and a source code, and these at least.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram showing the configuration of the face extractor concerning 1 operation gestalt of this invention.

[Drawing 2] It is drawing showing an example of a configuration of being used for the edge extract section 1.

[Drawing 3] It is drawing explaining an example of the edge image extracted in the edge extract section 1.

[Drawing 4] It is drawing showing an example of the template memorized by the template storage section 2.

[Drawing 5] It is the flow chart which shows the procedure of the vote processing performed in the vote section 4.

[Drawing 6] It is drawing explaining the concept of the number of votes memorized by vote processing in the vote storage region of the vote result storage section 3.

[Drawing 7] It is the flow chart which shows the procedure of the analysis processing performed in the analysis section 5.

[Drawing 8] It is drawing explaining the concept of the clustering processing performed at steps S23 and S24 of drawing 7.

[Drawing 9] It is drawing showing an example of the image edit processing performed in the image editorial department 6.

[Description of Notations]

1 — Edge extract section

2 — Template storage section

3 — Vote result storage section

4 — Vote section

5 — Analysis section

6 — Image editorial department

11 — Filter

12 — difference — the extract section

T1 — T3 — Template

t1—tn — Circle configuration

[Translation done.]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-222719

(P2001-222719A)

(43)公開日 平成13年8月17日(2001.8.17)

(51)Int.Cl.
G 0 6 T 7/60

2 5 0
1/00 3 4 0
7/00 3 0 0

識別記号
1 5 0

2 5 0
3 4 0
3 0 0

F I
G 0 6 T 7/60

1/00
7/00

テマコード(参考)
1 5 0 B 5 B 0 5 7
1 5 0 J 5 L 0 9 6

2 5 0 A
3 4 0 A
3 0 0 D

審査請求 未請求 請求項の数26 OL (全 12 頁) 最終頁に続く

(21)出願番号 特願2000-363449(P2000-363449)

(22)出願日 平成12年11月29日(2000.11.29)

(31)優先権主張番号 特願平11-342025

(32)優先日 平成11年12月1日(1999.12.1)

(33)優先権主張国 日本 (J P)

(71)出願人 000005821

松下電器産業株式会社
大阪府門真市大字門真1006番地

(72)発明者 松尾 英明

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 今川 和幸

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74)代理人 100098291

弁理士 小笠原 史朗

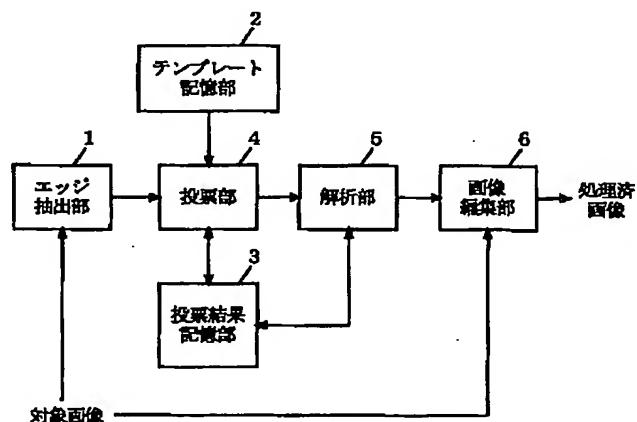
最終頁に続く

(54)【発明の名称】顔抽出装置及び顔抽出方法並びに顔抽出プログラムの記録媒体

(57)【要約】

【課題】 様々な対象画像から高速に顔の位置及び大きさを抽出できる、汎用性が高い顔抽出装置及び顔抽出方法を提供する。

【解決手段】 エッジ抽出部1は、対象画像からエッジ部を抽出してエッジ画像を生成する。テンプレート記憶部2には、予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートが記憶されている。投票結果記憶部3には、投票部4で行われる投票処理の結果を記憶する投票記憶領域が、テンプレートを構成する各サイズの形状毎に設けられている。投票部4は、エッジ画像の各画素位置において、テンプレートを用いて投票処理を行い、その投票結果を各投票記憶領域に記憶する。解析部5は、投票部4での投票処理が完了した後、各投票記憶領域に記憶された投票結果に基づき、そのクラスタを評価して、対象画像に含まれる顔の位置及び大きさを求める。



(2)

【特許請求の範囲】

【請求項 1】 対象画像に含まれる顔の位置及び大きさを求める顔抽出装置であって、前記対象画像からエッジ部（人物の外郭や顔の輪郭等に相当する画素）を抽出して、当該エッジ部だけの画像（以下、エッジ画像という）を生成するエッジ抽出部と、

予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶するテンプレート記憶部と、

前記テンプレートを構成する各サイズの形状毎に、前記エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する投票結果記憶部と、

前記エッジ部の各画素位置に前記テンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、前記各サイズの形状を形成する全画素の位置に対応する各座標位置について、前記投票結果記憶部に記憶されている前記投票数をそれぞれ増加又は減少させる投票部と、

前記投票結果記憶部に記憶されている前記各投票数に基づいて、前記対象画像に含まれる顔の位置及び大きさを求める解析部とを備える、顔抽出装置。

【請求項 2】 前記予め定めた形状は、円であることを特徴とする、請求項 1 に記載の顔抽出装置。

【請求項 3】 前記予め定めた形状は、楕円であること を特徴とする、請求項 1 に記載の顔抽出装置。

【請求項 4】 前記予め定めた形状は、多角形であることを特徴とする、請求項 1 に記載の顔抽出装置。

【請求項 5】 前記エッジ抽出部は、前記対象画像にフィルタリングを施して高周波成分を求めるこことにより、前記エッジ部を抽出することを特徴とする、請求項 1～4 のいずれかに記載の顔抽出装置。

【請求項 6】 前記エッジ抽出部は、前記対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、前記エッジ部を抽出することを特徴とする、請求項 1～4 のいずれかに記載の顔抽出装置。

【請求項 7】 前記エッジ抽出部は、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけを前記エッジ部とすることを特徴とする、請求項 5 又は 6 に記載の顔抽出装置。

【請求項 8】 前記解析部は、前記投票結果記憶部に記憶されている前記各投票数に対してクラスタリングを行い、前記対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする、請求項 1～7 のいずれかに記載の顔抽出装置。

【請求項 9】 前記解析部で求められた顔の位置及び大きさによって定められる顔領域と、それ以外の領域とを区別して、前記対象画像に所定の編集を施す画像編集部をさらに備える、請求項 1～8 のいずれかに記載の顔抽

3

出装置。

【請求項 10】 前記解析部で求められた顔の位置及び大きさによって定められる顔領域の画像を、他の画像で置き換える編集を施す画像編集部をさらに備える、請求項 1～8 のいずれかに記載の顔抽出装置。

【請求項 11】 対象画像に含まれる顔の位置及び大きさを求める顔抽出方法であって、

前記対象画像からエッジ部（人物の外郭や顔の輪郭等に相当する画素）を抽出して、当該エッジ部だけの画像（以下、エッジ画像という）を生成する抽出ステップと、

予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶する第 1 の記憶ステップと、

前記テンプレートを構成する各サイズの形状毎に、前記エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する第 2 の記憶ステップと、

前記エッジ部の各画素位置に前記テンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、前記各サイズの形状を形成する全画素の位置に対応する各座標位置について、前記第 2 の記憶ステップで記憶される前記投票数をそれぞれ増加又は減少させる投票ステップと、

前記投票ステップの処理後の前記各投票数に基づいて、前記対象画像に含まれる顔の位置及び大きさを求める解析ステップとを備える、顔抽出方法。

【請求項 12】 前記予め定めた形状は、円であることを特徴とする、請求項 11 に記載の顔抽出方法。

【請求項 13】 前記予め定めた形状は、楕円であることを特徴とする、請求項 11 に記載の顔抽出方法。

【請求項 14】 前記予め定めた形状は、多角形であることを特徴とする、請求項 11 に記載の顔抽出方法。

【請求項 15】 前記抽出ステップは、前記対象画像にフィルタリングを施して高周波成分を求めるこことにより、前記エッジ部を抽出することを特徴とする、請求項 11～14 のいずれかに記載の顔抽出方法。

【請求項 16】 前記抽出ステップは、前記対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、前記エッジ部を抽出することを特徴とする、請求項 11～14 のいずれかに記載の顔抽出方法。

【請求項 17】 前記抽出ステップは、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけを前記エッジ部とすることを特徴とする、請求項 15 又は 16 に記載の顔抽出方法。

【請求項 18】 前記解析ステップは、前記投票ステップの処理後の前記各投票数に対してクラスタリングを行い、前記対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする、請求項 11～17 のいず

(3)

れかに記載の顔抽出方法。

【請求項19】 対象画像に含まれる顔の位置及び大きさを求める顔抽出方法が、コンピュータ装置上で実行可能なプログラムとして記録された媒体であって、前記対象画像からエッジ部（人物の外郭や顔の輪郭等に相当する画素）を抽出して、当該エッジ部だけの画像（以下、エッジ画像という）を生成する抽出ステップと、

予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶する第1の記憶ステップと、

前記テンプレートを構成する各サイズの形状毎に、前記エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する第2の記憶ステップと、

前記エッジ部の各画素位置に前記テンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、前記各サイズの形状を形成する全画素の位置に対応する各座標位置について、前記第2の記憶ステップで記憶される前記投票数をそれぞれ増加又は減少させる投票ステップと、

前記投票ステップの処理後の前記各投票数に基づいて、前記対象画像に含まれる顔の位置及び大きさを求める解析ステップとを、少なくとも実行するためのプログラムを記録した、記録媒体。

【請求項20】 前記予め定めた形状は、円であることを特徴とする、請求項19に記載の記録媒体。

【請求項21】 前記予め定めた形状は、楕円であることを特徴とする、請求項19に記載の記録媒体。

【請求項22】 前記予め定めた形状は、多角形であることを特徴とする、請求項19に記載の記録媒体。

【請求項23】 前記抽出ステップは、前記対象画像にフィルタリングを施して高周波成分を求ることにより、前記エッジ部を抽出することを特徴とする、請求項19～22のいずれかに記載の記録媒体。

【請求項24】 前記抽出ステップは、前記対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、前記エッジ部を抽出することを特徴とする、請求項19～22のいずれかに記載の記録媒体。

【請求項25】 前記抽出ステップは、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけを前記エッジ部とすることを特徴とする、請求項23又は24に記載の記録媒体。

【請求項26】 前記解析ステップは、前記投票ステップの処理後の前記各投票数に対してクラスタリングを行い、前記対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする、請求項19～25のいずれかに記載の記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、顔抽出装置及び顔抽出方法並びに顔抽出プログラムの記録媒体に関し、より特定的には、画像処理に用いられる技術であって、対象画像に含まれる顔の位置及び大きさを、テンプレートを用いて高速に抽出する装置及び方法並びに当該方法を実行するためのプログラムを記録した媒体に関する。

【0002】

【従来の技術】言うまでもなく、顔は、人物の思考や感情を表すという重要な意味を持つ。従って、人物を含む画像を取り扱う画像処理の分野においては、静止画像や動画像（人物が実際に撮影された画像の他、コンピュータグラフィックス等で人為的に生成された画像も含む）について、人物の顔が対象画像のどの位置にどれくらいの大きさで存在しているかを、自動的に検出・処理できるシステムがあれば便利である。このため、近年、画像処理システム上で、対象となる画像から顔の領域を抽出しようとする試みが始まっている。

【0003】このような画像から顔の領域を抽出する従来の技術として、特開平9-73544号公報（以下、第1の文献という）や特開平10-307923号公報（以下、第2の文献という）等に開示されているものが存在する。この第1の文献には、顔領域を楕円で近似する技術が開示されており、楕円を規定する5つのパラメータ（中心座標（x, y）、半径r、長軸と短軸との比b及び長軸とx軸とのなす角θ）を逐次変更して、顔領域の抽出に最適な値を求めるを行っている。また、第2の文献には、顔の部品（目、鼻、口など）を順番に見つけて行く技術が開示されている。

【0004】

【発明が解決しようとする課題】しかしながら、上記第1の文献に記載された従来の技術では、各パラメータを変化させつつ近似の評価を行うことになるので、膨大な繰り返し計算が必要になる（特に、角θの処理が重いためである）。このため、現状のパーソナルコンピュータレベルの処理能力では、この従来の技術を用いて、変化する人物画像に追従させて実時間で顔抽出処理を行うことは、絶望的である。また、この従来の技術では、1つの画像に複数の人物の顔が含まれているような場合に対する配慮がなく、汎用性が低いと言わざるを得ない。また、上記第2の文献に記載された従来の技術では、まず、顔領域の位置が既知でなければ、顔の部品の抽出処理ができない。従って、この従来の技術では、特定の画像にしか適用することができず、汎用性が低い。

【0005】それ故、本発明の目的は、様々な対象画像から高速に顔の位置及び大きさを抽出できる、汎用性が高い顔抽出装置及び顔抽出方法並びに当該方法を実行するためのプログラムを記録した媒体を提供することである。

【0006】

(4)

5

【課題を解決するための手段および発明の効果】第1の発明は、対象画像に含まれる顔の位置及び大きさを求める顔抽出装置であって、対象画像からエッジ部（人物の外郭や顔の輪郭等に相当する画素）を抽出して、当該エッジ部だけの画像（エッジ画像）を生成するエッジ抽出部と、予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶するテンプレート記憶部と、テンプレートを構成する各サイズの形状毎に、エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する投票結果記憶部と、エッジ部の各画素位置にテンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、各サイズの形状を形成する全画素の位置に対応する各座標位置について、投票結果記憶部に記憶されている投票数をそれぞれ増加又は減少させる投票部と、投票結果記憶部に記憶されている各投票数に基づいて、対象画像に含まれる顔の位置及び大きさを求める解析部とを備える。上記のように、第1の発明によれば、処理負担が軽い投票処理とその評価だけで、顔の位置を高速に検出できる。しかも、相似で同心状の複数サイズの形状を備えたテンプレートを用いているので、顔領域を含むであろうエッジ部が、これらの形状のいずれのサイズに近いかという実質的な近似を行っていることになり、顔の大きさも高速に抽出できる。このように、第1の発明では、処理負担を大幅に軽減できるので、現状のパーソナルコンピュータレベルの処理能力でも、ほぼ実時間で顔領域を抽出することができる。また、第1の発明では、対象画像の内どの部分に顔領域があるかという点や顔の個数等は、抽出処理前に不明であって差し支えなく、広い範囲の様々な対象画像について一様に顔領域を検出できるので、極めて汎用性が高い。

【0007】第2～第4の発明は、それそれ第1の発明に従属する発明であって、予め定めた形状が、円、橢円又は多角形のいずれかであることを特徴とする。上記のように、第2～第4の発明は、テンプレートに用いる典型的な形状を示したものである。特に、中心点から形状を形成する各画素までの距離が一定である円をテンプレートの形状に用いることにより、投票結果の精度を高めさせることができる。

【0008】第5の発明は、第1～第4の発明に従属する発明であって、エッジ抽出部は、対象画像にfiltratingを施して高周波成分を求めるこにより、エッジ部を抽出することを特徴とする。上記のように、第5の発明によれば、対象画像にfiltratingを施して高周波成分を求めるので、対象画像が静止画像である場合における顔の位置及び大きさの検出に好適である。

【0009】第6の発明は、第1～第4の発明に従属する発明であって、エッジ抽出部は、対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、エッジ部を抽出することを特徴と

50

6

する。上記のように、第6の発明によれば、現在の対象画像とそれに時間的に前後する画像との差分を求めるので、対象画像が動画像である場合における、各コマでの顔の位置及び大きさの検出に好適である。また、検出にテンプレートを用いているので、ズームアップ時やパン時のように、顔の変化が大きい時でも安定して高速に顔領域を抽出できる。

【0010】第7の発明は、第5及び第6の発明に従属する発明であって、エッジ抽出部は、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけをエッジ部とすることを特徴とする。上記のように、第7の発明によれば、内側のテクスチャの差が生じる部分を排除した、人物の外郭のみからなるエッジ部を得ることができるので、顔領域の抽出処理を高速に行うことができる。

【0011】第8の発明は、第1～第7の発明に従属する発明であって、解析部は、投票結果記憶部に記憶されている各投票数に対してクラスタリングを行い、対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする。上記のように、第8の発明によれば、投票結果（各投票数）をクラスタとして取り扱うことにより、対象画像内に複数の顔が存在する場合でも、投票結果の相互の関連性を正当に評価して、顔領域を抽出することができる。

【0012】第9の発明は、第1～第8の発明に従属する発明であって、解析部で求められた顔の位置及び大きさによって定められる顔領域と、それ以外の領域とを区別して、対象画像に所定の編集を施す画像編集部をさらに備える。上記のように、第9の発明によれば、対象画像の顔の位置及び大きさによって定められる顔領域とそれ以外の領域とを、区別して編集することにより、顔のみを強調して、所望する部分が見易い画像を得ることができる。例えば、顔領域以外の部分を一定のテクスチャや色でベタ塗りにして、おもしろい効果を狙うことができる。

【0013】第10の発明は、第1～第8の発明に従属する発明であって、解析部で求められた顔の位置及び大きさによって定められる顔領域の画像を、他の画像で置き換える編集を施す画像編集部をさらに備える。上記のように、第10の発明によれば、対象画像の顔の位置及び大きさによって定められる顔領域の画像を、他の画像で置き換えることにより、顔を意図的に隠すことができる。例えば、痴呆性老人の徘徊を監視する動画像に用いる場合のように、顔の存在自体は十分に監視する必要があるが、プライバシーを保護すべき時、顔の部分を別の画像で置き換えて対応することができる。さらには、人間の動作を撮像して、これを他のキャラクタに置き換えるような場合にも、応用することができる。

【0014】第11の発明は、対象画像に含まれる顔の

(5)

7

位置及び大きさを求める顔抽出方法であって、対象画像からエッジ部を抽出して、エッジ画像を生成する抽出ステップと、予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶する第1の記憶ステップと、テンプレートを構成する各サイズの形状毎に、エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する第2の記憶ステップと、エッジ部の各画素位置にテンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、各サイズの形状を形成する全画素の位置に対応する各座標位置について、第2の記憶ステップで記憶される投票数をそれぞれ増加又は減少させる投票ステップと、投票ステップの処理後の各投票数に基づいて、対象画像に含まれる顔の位置及び大きさを求める解析ステップとを備える。上記のように、第1の発明によれば、処理負担が軽い投票処理とその評価だけで、顔の位置を高速に検出できる。しかも、相似で同心状の複数サイズの形状を備えたテンプレートを用いているので、顔領域を含むであろうエッジ部が、これらの形状のいずれのサイズに近いかという実質的な近似を行っていることになり、顔の大きさも高速に抽出できる。このように、第1の発明では、処理負担を大幅に軽減できるので、現状のパソコン用コンピュータレベルの処理能力でも、ほぼ実時間で顔領域を抽出することができる。また、第1の発明では、対象画像の内どの部分に顔領域があるかという点や顔の個数等は、抽出処理前に不明であって差し支えなく、広い範囲の様々な対象画像について一様に顔を検出できるので、極めて汎用性が高い。

【0015】第12～第14の発明は、それぞれ第11の発明に従属する発明であって、予め定めた形状が、円、楕円又は多角形のいずれかであることを特徴とする。上記のように、第12～第14の発明は、テンプレートに用いる典型的な形状を示したものである。特に、中心点から形状を形成する各画素までの距離が一定である円をテンプレートの形状に用いることにより、投票結果の精度を高くさせることができる。

【0016】第15の発明は、第11～第14の発明に従属する発明であって、抽出ステップは、対象画像にフィルタリングを施して高周波成分を求ることにより、エッジ部を抽出することを特徴とする。上記のように、第15の発明によれば、対象画像にフィルタリングをして高周波成分を求めるので、対象画像が静止画像である場合における顔の位置及び大きさの検出に好適である。

【0017】第16の発明は、第11～第14の発明に従属する発明であって、抽出ステップは、対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、エッジ部を抽出することを特徴とする。上記のように、第16の発明によれば、現在の対象画像とそれに時間的に前後する画像との差分を

求めるので、対象画像が動画像である場合における、各コマでの顔の位置及び大きさの検出に好適である。また、検出にテンプレートを用いているので、ズームアップ時やバーン時のように、顔の変化が大きい時でも安定して高速に顔領域を抽出できる。

【0018】第17の発明は、第15及び第16の発明に従属する発明であって、抽出ステップは、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけをエッジ部とすることを特徴とする。上記のように、第17の発明によれば、内側のテクスチャの差が生じる部分を排除した、人物の外郭のみからなるエッジ部を得ることができるので、顔領域の抽出処理を高速に行うことができる。

【0019】第18の発明は、第11～第17の発明に従属する発明であって、解析ステップは、投票ステップの処理後の各投票数に対してクラスタリングを行い、対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする。上記のように、第18の発明によれば、投票結果（各投票数）をクラスタとして取り扱うことにより、対象画像内に複数の顔が存在する場合でも、投票結果の相互の関連性を正当に評価して、顔領域を抽出することができる。

【0020】第19の発明は、対象画像に含まれる顔の位置及び大きさを求める顔抽出方法が、コンピュータ装置上で実行可能なプログラムとして記録された媒体であって、対象画像からエッジ部を抽出して、エッジ画像を生成する抽出ステップと、予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートを記憶する第1の記憶ステップと、テンプレートを構成する各サイズの形状毎に、エッジ画像上の座標位置と投票数とを対応付けてそれぞれ記憶する第2の記憶ステップと、エッジ部の各画素位置にテンプレートの中心点を順次移動させ、移動させた当該画素位置毎に、各サイズの形状を形成する全画素の位置に対応する各座標位置について、第2の記憶ステップで記憶される投票数をそれぞれ増加又は減少させる投票ステップと、投票ステップの処理後の各投票数に基づいて、対象画像に含まれる顔の位置及び大きさを求める解析ステップとを、少なくとも実行するためのプログラムを記録している。

【0021】第20～第22の発明は、それ第19の発明に従属する発明であって、予め定めた形状が、円、楕円又は多角形のいずれかであることを特徴とする。

【0022】第23の発明は、第19～第22の発明に従属する発明であって、抽出ステップは、対象画像にフィルタリングを施して高周波成分を求ることにより、エッジ部を抽出することを特徴とする。

【0023】第24の発明は、第19～第22の発明に

(6)

9

従属する発明であって、抽出ステップは、対象画像を構成する画像毎に、現在の画像と時間的に前後する画像との差分を求ることにより、エッジ部を抽出することを特徴とする。

【0024】第25の発明は、第23及び第24の発明に従属する発明であって、抽出ステップは、抽出した画素のうち、所定の矩形領域毎に、当該領域内の各走査ラインの最も左側及び右側に存在する画素を検出し、当該検出した画素だけをエッジ部とすることを特徴とする。

【0025】第26の発明は、第19～第25の発明に従属する発明であって、解析ステップは、投票ステップの処理後の各投票数に対してクラスタリングを行い、対象画像に含まれる顔の位置及び大きさの絞り込みを行うことを特徴とする。

【0026】

【発明の実施の形態】図1は、本発明の一実施形態に係る顔抽出装置の構成を示すブロック図である。図1において、本実施形態に係る顔抽出装置は、エッジ抽出部1と、テンプレート記憶部2と、投票結果記憶部3と、投票部4と、解析部5と、画像編集部6とを備える。以下、図面を参照しながら、本実施形態に係る顔抽出装置の各構成の動作及び顔抽出方法を説明する。

【0027】エッジ抽出部1は、顔抽出処理の対象となる画像（以下、対象画像という）を入力し、この対象画像からエッジ部を抽出してエッジ部だけの画像（以下、エッジ画像という）を生成する。ここで、エッジ部とは、人物の外郭や顔の輪郭等に相当する部分（画素）であって、対象画像内の高周波成分となる部分である。入力される対象画像は、静止画像又は動画像のいずれでもよく、エッジ抽出部1は、それぞれ次の手法を用いて、対象画像からエッジ部を抽出することができる。

【0028】静止画像を対象画像とする場合には、図2(a)に示すように、エッジ抽出部1として、高周波成分を取り出すフィルタ11を用いる。これにより、対象画像をフィルタ11に通すだけで、エッジ部を抽出することができる。このフィルタ11としては、Sobelフィルタが好適に用いられる。動画像を対象画像とする場合には、図2(b)に示すように、エッジ抽出部1として、現在の画像と時間的に前後する画像との差分（画素毎のデータ差分）を抽出する差分抽出部12を用いる。これにより、動画像を構成する画像毎に差分をそれ求め、この差分が大きい部分（動きが生じた部分）を、エッジ部として抽出することができる。

【0029】なお、上述した手法では、人物の外郭や顔の輪郭等と共に、その内側においてテクスチャの差が生じる部分も、エッジ部として抽出される。そのようなエッジ部が抽出されたエッジ画像の一例を、図3(a)に示す。本発明の顔抽出装置においては、テクスチャの差が生じる部分がエッジ部として抽出されても、問題なく処理できるが、処理をより高速に行えるようにするた

(6)

10

め、次のような手法を用いることが好ましい。まず、図3(a)のようなエッジ画像において、ある程度集中して存在するエッジ部の範囲が、矩形領域でそれぞれ囲まれる（図3(b)）。次に、この各矩形領域内で左端及び右端の双方から走査がそれぞれ行われ、最初に検出されたエッジ部のみが対象画像のエッジ部として決定される（図3(c)）。この処理を矩形領域内の全走査ラインについて行うことにより、図3(c)に示すように、内側のテクスチャの差が生じる部分を排除した、人物の外郭のみからなるエッジ部を得ることができる。このような処理を行う構成は、上述したフィルタ11又は差分抽出部12の後段に挿入されればよい。

【0030】テンプレート記憶部2には、予め定めた形状を、相似で大きさを異ならせた種々のサイズによって、中心点で同心状に複数設けたテンプレートのデータが記憶されている。このテンプレートの形状には、円、橢円、正多角形、多角形等を用いることができるが、中心点から形状線（形状を形成する各画素）までの距離が常に一定である円を用いることが最も好ましい。これにより、後述する投票結果の精度を高くさせることができる。以下、本実施形態では、図4に示すように、中心点がPで半径が異なる同心円を複数設けたテンプレートを用いた場合を説明する。ここで、テンプレートを構成する複数の円t1～tn(nは、任意の整数)は、図4(a)に示すテンプレートT1のように、一定間隔で半径が変化する構成であってもよいし、図4(b)に示すテンプレートT2のように、不定間隔で半径が変化する構成であってもよい。また、テンプレートを構成する複数の円t1～tnは、図4(b)に示すテンプレートT2のように、全ての線幅が1ドット（対象画像の1画素に相当）で構成されてもよいし、図4(c)に示すテンプレートT3のように、一部又は全部の線幅が2ドット以上（すなわち、円環形状）で構成されてもよい。なお、以下の説明では、円及び円環を総称して単に「円」という。

【0031】この複数の円t1～tnは、まとめて1つのテンプレートとして扱われてテンプレート記憶部2に記憶されるが、実際の処理では、テンプレートを構成する各円t1～tnは、独立して扱われることとなる。このため、各円t1～tnを形成する画素データは、テンプレート記憶部2において、例えばテーブル形式でそれぞれ記憶される。

【0032】投票結果記憶部3には、後述する投票部4において行われる投票処理の結果を記憶する領域（以下、投票記憶領域という）が、テンプレート記憶部2に記憶されているテンプレートを構成する各サイズの形状毎に、設けられている。この例では、各サイズの形状が円t1～tnであるので、投票結果記憶部3には、円t1～tnに関してn個の投票記憶領域が設けられることとなる。なお、この投票記憶領域は、対象画像に対応す

(7)

11

る範囲を有する。

【0033】投票部4は、エッジ抽出部1で生成されたエッジ画像について、テンプレート記憶部2に記憶されているテンプレートを用いて、投票処理を行う。図5は、投票部4で行われる投票処理の手順を示すフローチャートである。図5を参照して、投票部4は、まず、投票結果記憶部3にアクセスして、各投票記憶領域内の座標を表す成分(投票値)を、全て零に初期化する(ステップS11)。次に、投票部4は、エッジ画像内のエッジ部の先頭画素位置に、テンプレートの中心点Pをセットする(ステップS12)。この先頭画素位置は、例えば、エッジ画像上を左上から右上又は左下へ順次走査して行き、最初に検出されたエッジ部の画素の位置とすればよい。

【0034】次に、投票部4は、テンプレートを構成する形状(この例では、円 $t_1 \sim t_n$)を特定するカウンタ*i*を、「1」に初期化する(ステップS13)。次に、投票部4は、カウンタ*i*(=1)によって特定される円 t_1 について、円 t_1 を形成する全画素のエッジ画像上のx y座標をそれぞれ取得する(ステップS14)。そして、投票部4は、投票結果記憶部3に設けられた円 t_1 に関する投票記憶領域において、取得した各x y座標を表す成分にそれぞれ「1」を加算して投票を行う(ステップS15)。この処理が終わると、投票部4は、カウンタ*i*を1つインクリメントして、*i*=2とする(ステップS17)。次に、投票部4は、カウンタ*i*(=2)によって特定される円 t_2 について、円 t_2 を形成する全画素のエッジ画像上のx y座標をそれぞれ取得する(ステップS14)。そして、投票部4は、投票結果記憶部3に設けられた円 t_2 に関する投票記憶領域において、取得した各x y座標を表す成分にそれぞれ「1」を加算して投票を行う(ステップS15)。

【0035】以降同様にして、投票部4は、*i*=*n*になるまでカウンタ*i*を1つずつインクリメントしながら(ステップS16, S17)、テンプレートを構成する全形状である円 $t_3 \sim t_n$ について、上記ステップS14及びS15の投票処理を繰り返し行う。これにより、各円 $t_1 \sim t_n$ に関する投票記憶領域のそれそれに、先頭画素位置における投票処理が行われることになる。そしてさらに、投票部4は、エッジ部の次の画素位置にテンプレートの中心点Pをセットして上記ステップS13～S17の処理を繰り返し行うことを、エッジ画像内のエッジ部の全画素に対して、1回ずつ行う(ステップS18, S19)。すなわち、投票部4による投票処理は、テンプレートの中心点Pがエッジ部の全画素を這うように行われる。

【0036】例えば、図3(c)に示すエッジ画像に上記投票処理を施すことによって、投票結果記憶部3に設けられた*n*個の投票記憶領域には、図6に示すような投票数が記憶される。なお、図6では、図面を見易くする

(7)

12

ため、エッジ部の一部の画素位置で投票処理が行われた場合を示している。図6において、実線円の部分が、上記ステップS15においてテンプレートの各サイズの形状(円 $t_1 \sim t_n$)に基づいて投票された座標成分に相当し、座標数「1」となる。また、上述したように各投票数は累積加算されるので、図6の実線円が交差する部分(図中、●印で示す)は、交差する数が多いほど投票数が高いことを表している。

【0037】そのため、中心点を持つ円又は楕円に近似した顔の輪郭を表現するエッジ部に、上述した投票処理を施せば、その中心点付近に高い投票数が集中することとなる。従って、高い投票値が集中する部分を判断すれば、顔の中心を特定することが可能になる。また、このような高い投票値が集中する現象は、テンプレートの中でも、顔の輪郭を表現するエッジ部の最小幅と等しい又は非常に近い半径を持つ円形状を用いた場合に、より顕著に現れる。従って、この現象がどの円形状の投票記憶領域に顕著に現れているかを判断すれば、顔の大きさを特定することが可能になる。この点は、一般化ハフ変換と似ていると言える。しかし、本発明の顔抽出方法では、同心状に複数サイズの形状を持つテンプレートを使用することにより、エッジ部の中心点と共にその大きさも一度に特定できるという点で、一般化ハフ変換とは明確に異なる。

【0038】なお、上記ステップS11において、各投票記憶領域内の座標を表す成分を、全て予め定めた最大値に初期化し、上記ステップS15において、取得した各x y座標を表す成分からそれぞれ「1」を減算して投票を行ってもよい。この場合、低い投票値が集中する部分を判断すれば、顔の中心を特定することが可能であり、この集中現象がどの円形状の投票記憶領域に顕著に現れているかを判断すれば、顔の大きさを特定することが可能になる。また、上記ステップS15において、投票数を加算又は減算させる値は「1」以外であってもよく、値を自由に設定することができる。

【0039】次に、投票結果記憶部3に記憶された投票結果に基づいて、対象画像の顔領域を特定する手法を説明する。解析部5は、投票部4による投票処理が完了した後、投票結果記憶部3に記憶された投票結果に基づいて、そのクラスタを評価して、対象画像に含まれる顔の位置及び大きさを求める。図7は、解析部5で行われる解析処理の手順を示すフローチャートである。

【0040】図7を参照して、解析部5は、まず、テンプレートを構成する形状(この例では、円 $t_1 \sim t_n$)を特定するカウンタ*j*を、「1」にセットする(ステップS21)。次に、解析部5は、カウンタ*j*(=1)によって特定される円 t_1 について、投票結果記憶部3の円 t_1 に関する投票記憶領域に記憶されている投票結果を参照して、投票数が予め定めたしきい値G(例えば、200等)を越える成分だけを抽出する(ステップS2

(8)

13

2)。このしきい値Gは、対象画像の精細度や所望する抽出精度に基づいて、任意に定めることができる。次に、解析部5は、抽出した成分だけを対象に、クラスタリングを行い(ステップS23)、クラスタ化された各領域の分散値及び共分散値をそれぞれ計算する(ステップS24)。このクラスタリングにおける類似度は、ユークリッド平方距離、標準化ユークリッド平方距離、マハラノビスの汎距離又はミンコフスキーダイアゴナル距離のいずれを用いて判断されてもよい。また、クラスタの形成には、最短距離法(SLINK:single linkage clustering method)、最長距離法(CLINK:complete linkage clustering method)又は群平均法(UPGMA:unweighted pair-group method using arithmetic averages)のいずれを用いててもよい。次に、解析部5は、クラスタ化された各領域の分散値及び共分散値を、予め定めたしきい値Hと比較する(ステップS25)。そして、このステップS25において各値がしきい値H未満の場合、解析部5は、その領域の中心点を顔の中心点とみなして、この時のカウンタj(=1)が指す円t1のサイズ(直径)を顔の短軸長とし(ステップS26)、この短軸長に一定値(経験的に定める)を加えた長さを顔の長軸長として決定する(ステップS27)。そして、解析部5は、この決定した中心点、短軸長及び長軸長を、解析結果として保持する(ステップS28)。一方、上記ステップS25において各値がしきい値H以上の場合は、解析部5は、その領域の中心点が顔の中心点ではないと判断して、次の処理に移る。

【0041】この処理が終わると、解析部5は、カウンタjを1つインクリメントして、j=2とする(ステップS30)。次に、解析部5は、カウンタj(=2)によって特定される円t2について、投票結果記憶部3の円t2に関する投票記憶領域に記憶されている投票結果を参照して、投票数が予め定めたしきい値Gを越える成分だけを抽出する(ステップS22)。次に、解析部5は、抽出した成分だけを対象に、クラスタリングを行い(ステップS23)、クラスタ化された各領域の分散値及び共分散値をそれぞれ計算する(ステップS24)。次に、解析部5は、クラスタ化された各領域の分散値及び共分散値を、予め定めたしきい値Hと比較する(ステップS25)。そして、このステップS25において各値がしきい値H未満の場合、解析部5は、その領域の中心点を顔の中心点とみなして、この時のカウンタj(=2)が指す円t2のサイズを顔の短軸長とし(ステップS26)、この短軸長に一定値を加えた長さを顔の長軸長として決定する(ステップS27)。そして、解析部5は、この決定した中心点、短軸長及び長軸長を、解析結果として追加して保持する(ステップS28)。一方、上記ステップS25において各値がしきい値H以上の場合は、解析部5は、その領域の中心点が顔の中心点ではないと判断して、次の処理に移る。

14

【0042】以降同様にして、解析部5は、j=nになるまでカウンタjを1つずつインクリメントしながら(ステップS29, S30)、投票結果記憶部3に記憶されている各円t3～tnに関する投票記憶領域について、上記ステップS22～S28の解析処理を繰り返し行う。これにより、各円t1～tnに関する投票記憶領域における、顔領域抽出の解析結果を得ることができ。この解析結果は、画像編集部6へ出力される。

【0043】ここで、上記ステップS23及びS24で行われるクラスタリング処理を、図8を参照して簡単に説明する。投票数がしきい値Gを越える成分(図中●印)が、図8(a)のように分布している場合を想定する。このような場合、解析部5では、以下のようなクラスタリングによる評価が行われる。初期のクラスタリングでは、例えば、図8(b)のように4つの初期クラスタA, B, C及びDが生成される。初期クラスタが生成されると、次に、初期クラスタA, B, C及びDの相互間の類似度が求められ、類似度が所定のしきい値以上であれば、そのクラスタ同士の結合が行われる。図8(c)の例では、クラスタCとクラスタDとが結合され、クラスタEとなっている場合である。そして、最終的なクラスタA, B及びEの分散値等が計算されて評価され、分散値が小さいクラスタA及びBの中心点が顔の中心とみなされる。クラスタEは、分散値が大きくなるので、顔の中心とはみなされない。

【0044】なお、分散値等の評価によって、2つ以上のクラスタが検出された場合、次の要領に基づいて顔領域が判断されるとよい。第1として、中心点が同一でかつ大きさが異なるクラスタが複数検出された場合は、これらのクラスタのうち最小の分散値を持つクラスタを顔領域とする。第2として、中心点が異なりかつ大きさが異なるクラスタが複数検出された場合は、違う場所に違う大きさの顔領域が存在するものと判断する。第3として、中心点が異なりかつ大きさが同一のクラスタが複数検出された場合は、違う場所に同じ大きさの顔領域が存在するものと判断する。

【0045】画像編集部6は、解析部5から解析結果(顔領域)を入力し、解析結果によって顔領域とそれ以外の領域とを区別できることを利用して、対象画像に対して所望する様々な処理を施す。例えば、画像編集部6は、図9(a)に示す対象画像を、図9(b)に示すように顔領域のみを残して背景を切り取るか、背景を1色又は単一のテクスチャでペタ塗りすることによって、顔のみを強調した画像を得ることができる。また、図9(a)に示す対象画像の顔領域の画像を、図9(c)に示すように他の画像(他のキャラクターの顔画像等)に置き換えることができる。これにより、顔を意図的に隠すことが可能となる。なお、この画像編集部6は、抽出された顔領域を用いて所望する様々な処理を施すために必要に応じて設けられる構成であって、本発明の顔抽出装

(9)

15

置に必須の構成ではない。

【0046】以上のように、本発明の一実施形態に係る顔抽出装置及び方法によれば、負担が軽い投票処理（基本的には加算処理のみ）と投票数の評価だけで、顔の位置を高速に検出できる。しかも、相似で同心状の複数サイズの形状を備えたテンプレートを用いているので、顔領域であろうエッジ部が、これらの形状のいずれのサイズに近いかという実質的な近似を行っていることになり、顔の大きさも高速に抽出できる。このように、本発明では、処理負担を大幅に軽減できるので、現状のパソコン用コンピュータレベルの処理能力でも、ほぼ実時間で顔領域を抽出することができる。また、本発明では、対象画像の内どの部分に顔領域があるかという点や顔の個数等は、抽出処理前に不明であって差し支えがなく、広い範囲の様々な対象画像について一様に顔を検出できるので、極めて汎用性が高い。特に、投票結果をクラスタリングすることにより、対象画像内に複数の顔が存在する場合でも、投票結果の相互の関連性を正当に評価して顔領域を抽出することができる。

【0047】なお、典型的には、上記実施形態に係る顔抽出装置が実現する各機能（顔抽出方法）は、所定のプログラムが格納された記憶装置（ROM, RAM, ハードディスク等）と、当該プログラムを実行するCPU（セントラル・プロセシング・ユニット）によって実現される。この場合、各プログラムは、CD-ROMやフロッピー（登録商標）ディスク等の記録媒体を介して導入されてもよい。この導入には、プログラムが複数の記録媒体に分散的に記録されて配布される場合を含む。

【0048】ここで、プログラムが、オペレーティングシステムの一部であるか否かを問わず、種々のプロセス又はスレッド（DLL等）にその機能の一部を肩代わりさせている場合には、肩代わりさせている機能に係るプログラム部分が格納されていない記録媒体であっても、本発明の顔抽出方法を実行するプログラムが記録された「記録媒体」に該当する。

【0049】また、本発明の顔抽出方法は、スタンドアロン形式（図1）で実現される場合を例示したが、サーバ／クライアント形式で実現されてもよい。つまり、1つの端末機のみに顔抽出方法の全ての機能が含まれるスタンドアロン形式の他に、クライアントである端末機に接続可能なサーバ又はネットワーク上に、顔抽出方法の全部又は一部の機能が含まれるサーバ／クライアント形式でも、差し支えない。例えば、機能のほとんどをサーバ側に持たせ、クライアント側では、WWWブラウザだけを持たせるようにする。この場合、各種の情報（テンプレートや投票数等）は、通常サーバ上にあり、基本的

16

にネットワークを経由してクライアントへ配布されるが、その情報がサーバ上にある時は、そのサーバの記憶装置が「記録媒体」に該当し、その情報がクライアント上にある時は、そのクライアントの記録装置が「記録媒体」に該当することとなる。

【0050】さらに、本発明の顔抽出方法を実行するプログラムは、次のいずれの形態であってもよい。コンパイルされて機械語になったアプリケーションの形態。上述のプロセス又はスレッドにより解釈される中間コードの形態。少なくともリソース及びソースコードと、これらから機械語のアプリケーションを生成できるコンパイラ及びリンクなどが、「記録媒体」に格納される形態。少なくともリソース及びソースコードと、これらから中間コードのアプリケーションを生成できるインタプリタとが、「記録媒体」に格納される形態。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る顔抽出装置の構成を示すブロック図である。

【図2】エッジ抽出部1に用いられる構成の一例を示す図である。

【図3】エッジ抽出部1で抽出されるエッジ画像の一例を説明する図である。

【図4】テンプレート記憶部2に記憶されているテンプレートの一例を示す図である。

【図5】投票部4で行われる投票処理の手順を示すフローチャートである。

【図6】投票処理によって投票結果記憶部3の投票記憶領域に記憶される投票数の概念を説明する図である。

【図7】解析部5で行われる解析処理の手順を示すフローチャートである。

【図8】図7のステップS2.3及びS2.4で行われるクラスタリング処理の概念を説明する図である。

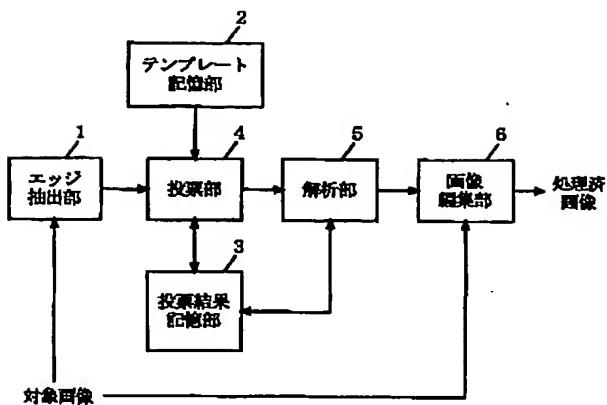
【図9】画像編集部6で行われる画像編集処理の一例を示す図である。

【符号の説明】

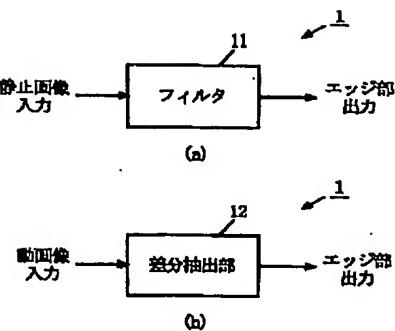
- 1…エッジ抽出部
- 2…テンプレート記憶部
- 3…投票結果記憶部
- 4…投票部
- 5…解析部
- 6…画像編集部
- 1 1…フィルタ
- 1 2…差分抽出部
- T 1～T 3…テンプレート
- t 1～t n…円形状

(10)

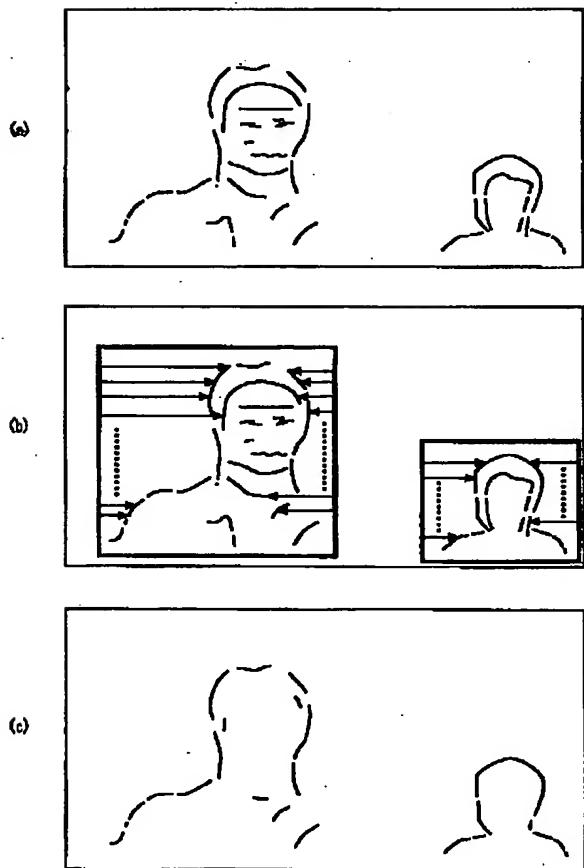
【図1】



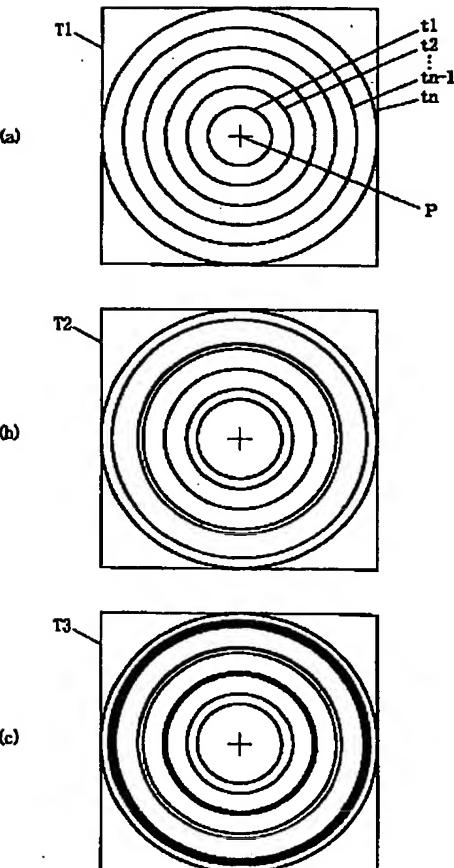
【図2】



【図3】

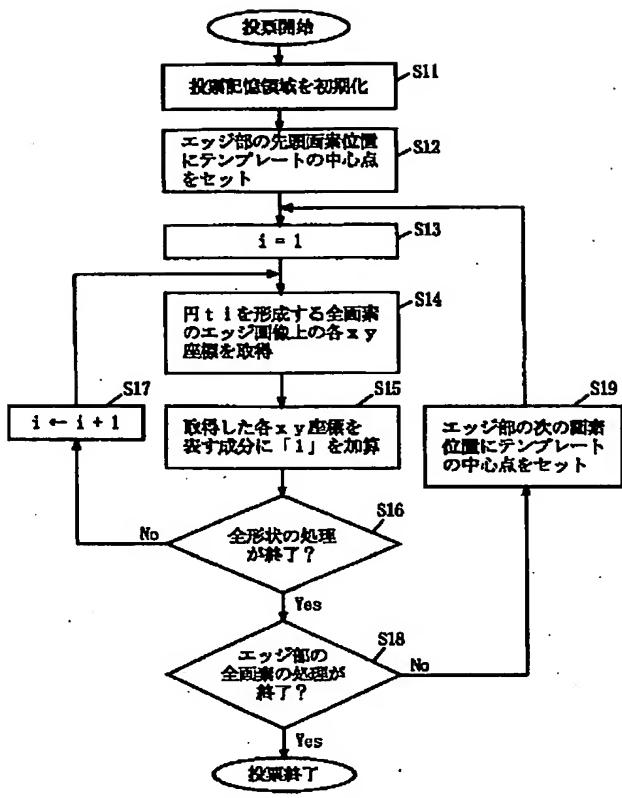


【図4】

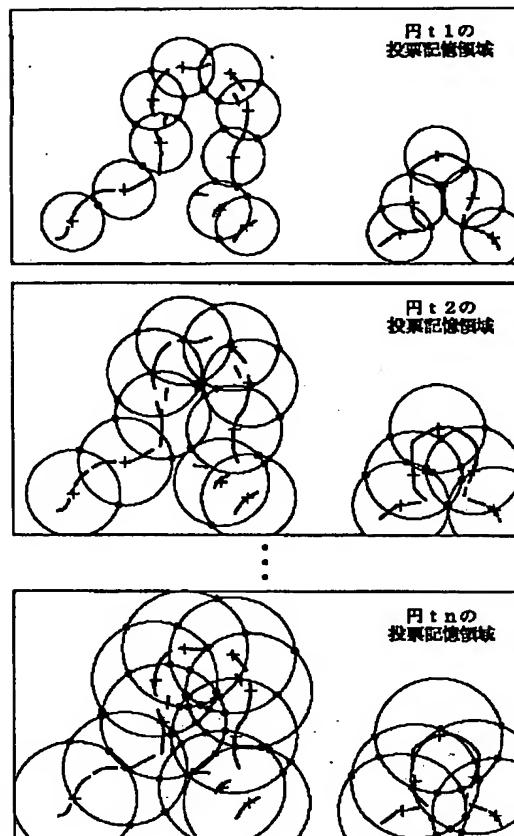


(11)

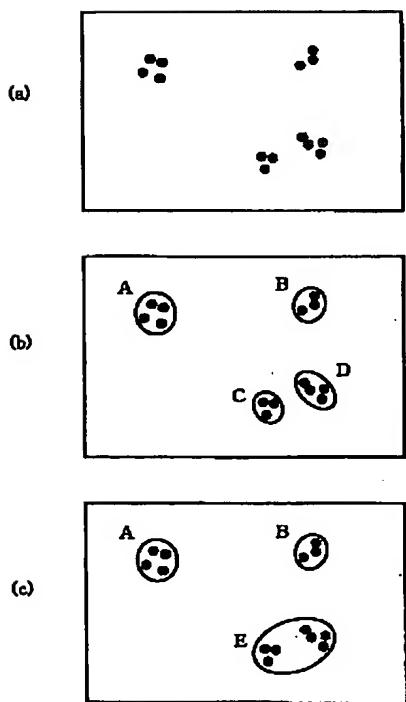
【図5】



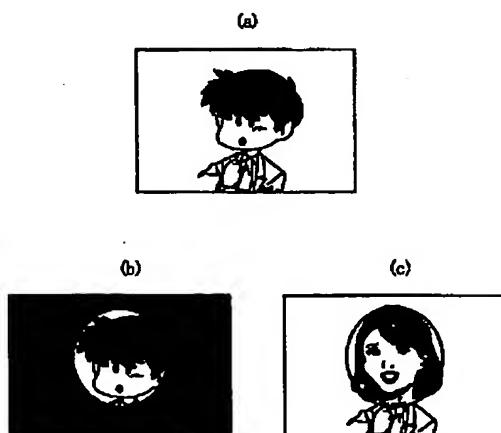
【図6】



【図8】



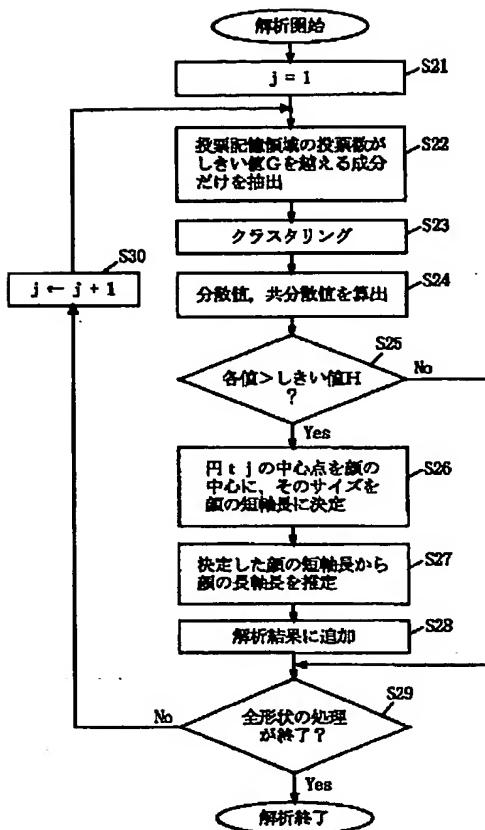
【図9】



BEST AVAILABLE COPY

(12)

【図7】



フロントページの続き

(51) Int. Cl. 7

G 06 T 7/20

識別記号

F I

G 06 T 7/20

テマコード (参考)

Z

(72) 発明者 高田 雄二

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 江島 俊朗

福岡県北九州市八幡西区千代3丁目7-7

(72) 発明者 馬▲場▼ 功淳

福岡県飯塚市横田753-23 IMビル202

F ターム (参考) 5B057 AA20 CA12 CA16 DA07 DB02

DC16

5L096 FA06 FA59 FA69 GA08 HA01

HA08 JA09